

Hopping for Success

SCAFFOLD “HOPPING” WITH CONTEXT

Tyler Peryea, Rajarshi Guha, John Braisted,
Noel Southall, Ajit Jadhav, Trung Nguyen

April 8, 2013

NCATS

Introduction

- Scaffold is a fundamental concept in medicinal chemistry
- Scaffold-based analyses are an integral part of early stage discovery
 - » Lead identification from high throughput screening (HTS) campaigns
 - » Lead optimization
 - » Lead “instrumentation,” i.e., addressing early stage liabilities such as IP and ADME/Tox.
- No freely available tools that meet our needs

Outline

- Overview
- Features of Scaffold Hopper
- A quick tour of Scaffold Hopper
- Future directions

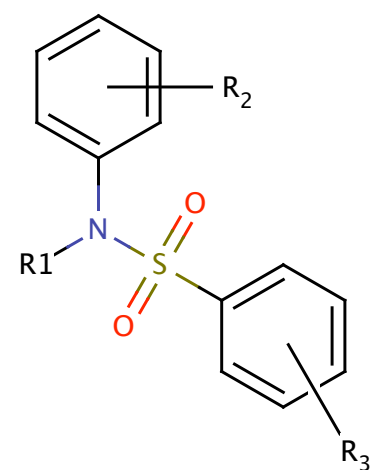
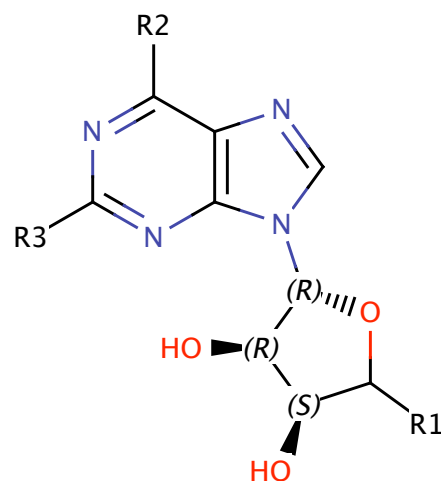
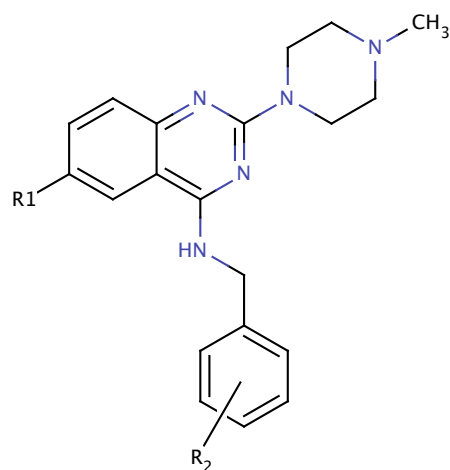
Overview

- Scaffold Hopper is a freely available tool that was initially developed in-house for automated R-group analysis
- Self-contained Java webstart application running inside the Java's secure sandbox
 - » Require explicit user's permissions for basic operations (e.g., file IO)
 - » Does not transmit user's structures over the network; communication with the server only over SSL.
 - » Heavy-lifting (e.g., computing MCS) is on the client side
- Look out for additional details on our blog
<http://tripod.nih.gov>
- Or take it out for a spin now
<https://tripod.nih.gov/ws/hopper/hopper.jnlp>

Software features

- Automatically generate “reasonable” R-group tables for a given dataset
- Scaffold-based “clustering” of the data
- “Bird’s eye” view of the data through scaffold network visualization
- Scaffold “hopping” (in the literal sense) in the context of publications (and soon targets and assays)
- PubMed on “steroid”
 - » Import data directly from PubMed ID or DOI
 - » Structure searching (sub-, super-, exact, and similarity) against PubMed
 - » Retrieve structures from PubMed’s text searches

Examples of scaffolds generated directly from Scaffold Hopper



A quick tour of Scaffold Hopper

Structure
searching

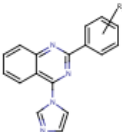
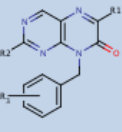
Project generated
scaffolds onto a new
dataset

Text searching

File Options Search:

Scaffolds (234) Network Singletons (2)

clk4-o1.sdf

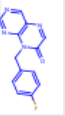
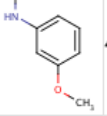
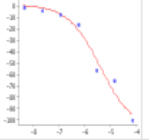
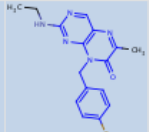
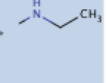

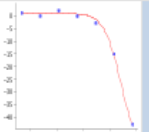
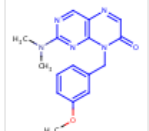
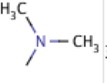
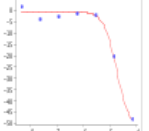
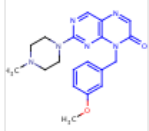
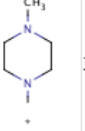
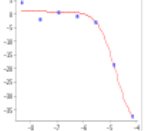
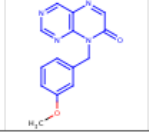
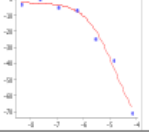
Scaffold	Score	Complexity	Count	clk4-o1-ptl2-3/set1-LogAC
	2.445	375	6	-5.1 ± 0.45
	4.346	312	121	-5.2 ± 1.05

Double-click to edit scaffold

Add additional custom scaffold

Add Sort Extend

References (1) Compounds (1354) R-group (121)

Structure	R1	R2	R3	clk4-o1-ptl2-3/set1-CRC	clk4-o1-ptl2-3/set1-AC50	PUBCHEM_SID
			4-flouro		3.981	4238063
	H ₃ C —				25.119	4237863
			3-methoxy		17.783	4239458
			3-methoxy		15.849	4239862
			3-methoxy		15.849	4240123

Quick tour (cont'd)

Entrez query syntax

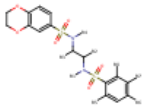
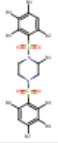
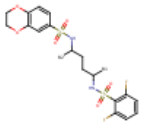
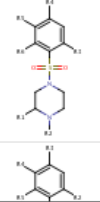
NCGC Scaffold Hopper

File Options

Search: min shen AND j med chem[journal]

Scaffolds (11) Network Singletons (1)

/chembl/v15/compounds/20017496/pubmed

Scaffold	Score	Complexity	Count	ALogP
	2.653	420	28	1.6 ± 0.51
	3.031	396	35	1.8 ± 0.56
	1.626	438	3	2.5 ± 0.30
	3.831	246	39	1.7 ± 0.64

Add Sort Extend

References (5) Compounds (58) R-group

[2010] Evaluation of substituted N,N'-diarylsulfonamides as activators of the tumor cell specific...

Title Evaluation of substituted N,N'-diarylsulfonamides as activators of the tumor cell specific M2 isoform of pyruvate kinase.

DOI 10.1021/jm901577g

PubMed [20017496](#)

Compounds 58

Abstract The metabolism of cancer cells is altered to support rapid proliferation. Pharmacological activators of a tumor cell specific pyruvate kinase isozyme (PKM2) may be an approach for altering the classic Warburg effect characteristic of aberrant metabolism in cancer cells yielding a novel antiproliferation strategy. In this manuscript, we detail the discovery of a series of substituted N,N'-diarylsulfonamides as activators of PKM2. The synthesis of numerous analogues and the evaluation of structure-activity relationships are presented as well as assessments of mechanism and selectivity. Several agents are found that have good potencies and appropriate solubility for use as chemical probes of PKM2 including 55 (AC(50) = 43 nM, maximum response = 84%; solubility = 7.3 microg/mL), 56 (AC(50) = 99 nM, maximum response = 84%; solubility = 5.7 microg/mL), and 58 (AC(50) = 38 nM, maximum response = 82%; solubility = 51.2 microg/mL). The small molecules described here represent first-in-class activators of PKM2.

Authors Boxer, MB.; Jiang, JK.; Vander Heiden, MG.; Shen, M.; Skoumbourdis, AP.; Southall, N.; Veith, H.; Leister, W.; Austin, CP.; Park, HW.; Ingles, J.; Cantley, LC.; Auld, DS.; Thomas, CJ.

Affiliation NIH Chemical Genomics Center, National Human Genome Research Institute, National Institutes of Health, 9800 Medical Center Drive, MSC 3370 Bethesda, Maryland 20850, USA.

Journal Journal of medicinal chemistry

[2010] Identification and optimization of inhibitors of Trypanosomal cysteine proteases: cruzain

Click on hyperlink to download compounds

Quick tour (cont'd)

NCGC Scaffold Hopper

File Options

Search:

Scaffolds (25) Network Singletons (0)

Click on scaffold and corresponding PubMed references are downloaded automatically

References (19) Compounds (300) R-group (163)

[2010] Development, validation, and use of quantitative struct...

Title Development, validation, and use of quantitative structure-activity relationship models of 5-hydroxytryptamine (2B) receptor ligands to identify novel receptor binders and putative valvulopathic compounds among common drugs.

DOI 10.1021/jm100600y

PubMed [20958049](#)

Compounds 12

Abstract Some antipsychotic drugs are known to cause valvular heart disease by activating serotonin 5-HT(2B) receptors. We have developed and validated binary classification QSAR models capable of predicting potential 5-HT(2B) actives. The classification accuracies of the models built to discriminate 5-HT(2B) actives from the inactives were as high as 80% for the external test set. These models were used to screen in silico 59,000 compounds included in the World Drug Index, and 122 compounds were predicted as actives with high confidence. Ten of them were tested in radioligand binding assays and nine were found active, suggesting a success rate of 90%. All validated actives were then tested in functional assays, and one compound was identified as a true 5-HT(2B) agonist. We suggest that the QSAR models developed in this study could be used as reliable predictors to flag drug candidates that are likely to cause valvulopathy.

Authors Hajjo, R.; Grulke, CM.; Golbraikh, A.; Setola, V.; Huang, XP.; Roth, BL.; Tropsha, A.

Future directions

- Integration with BARD
- Incorporate additional contexts; e.g., targets, assays, clinical trials, drugs, patents, etc.
- Explore interesting use cases, e.g.,
 - » Given a set of hits from a phenotypic screen, can we identify likely targets and/or pathways?

Acknowledgements

- Deepak Bandyopadhyay (GSK)
- Min Shen (NCGC)